The goal of our study is to describe a quantitative software-based semi-automated method to characterize Hoffa-synovitis.
**Methods:** Forty subjects were selected from the Osteoarthritis Initiative (OAI), a multicenter cohort of 4796 participants with or at risk for knee osteoarthritis (OA). Sagittal 3T TSE intermediated-weighted fat-suppressed (Iw FS) MRI of the knee were evaluated. A software method was used to characterize the regions of infrapatellar Hoffa's fat pad edema (surrogate for synovitis) on each slice. As an initial step, a center slice was defined as the location of the midportion of the ACL. The reader drew a region of interest delineating Hoffa's fat pad using an average of 8 slices medial and lateral to the patellar tendon to include the anatomy of the infrapatellar Hoffa's fat pad. The software automatically applied a quantitative algorithm on each image to detect signal associated with Hoffa-synovitis and produce an aggrate Software score for the entire scan. The software method was compared to the MRI Osteoarthritis Knee Score (MOAKS) Hoffa-synovitis score. The correspondence between the MOAKS score and the quantitative assessment was measuring using Spearman's rank correlation.
**Results:** Patients had an average age of 65.8 years at baseline, and 50% were female, with an average BMI of 29.0 at baseline. Following the short reader training, the software method was efficient, requiring less than 5 minutes per knee of reader time. Figure 1 provides a graph of the average Software score as a function of the MOAKS grade. The MOAKS score were distributed as follows: 0: $n = 5$, 1: $n = 22$, 2: $n = 13$. The quantitative synovitis measurement correlated moderately with MOAKS synovitis scores ($r = .51$). An ANOVA used to test for differences in mean measurements by MOAKS level was significant ($p = .0008$). Using the Tukey method, pair-wise comparisons all levels except 0 and 1 were significant at $p < .05$.
**Conclusions:** To our knowledge, this is the first study that describes a fully quantitative software tool to quantify Hoffa-synovitis. This method can potentially increase objectivity, accuracy and responsiveness. Once the measurement is fully validated, it will be feasible to provide a measurement for a large number of knees.

## 464
## CROSS-SECTIONAL AND LONGITUDINAL RELIABILITY OF SEMIQUANTITATIVE OSTEOARTHRITIS ASSESSMENT AT 1.0T EXTREMITY MRI: MULTI-READER DATA FROM THE MOST STUDY

F.W. Roemer †, M. Nevitt ‡, D. Felson †, M.D. Crema †, M.D. Marra †, J. Niu †, J.A. Lynch ‡, I. Tolstykh ‡, C. Lewis §, J. Torner ‖, A. Guermazi †. † *Boston Univ. Sch. of Med., Boston, MA;* ‡ *UCSF, San Francisco, CA;* § *Univ. of Alabama at Birmingham, Birmingham, AL;* ‖ *Univ. of Iowa, Iowa City, IA*

**Purpose:** Several large epidemiologic osteoarthritis (OA) studies including magnetic resonance imaging (MRI) are currently ongoing. A large proportion of these MRI datasets is being assessed in semi-quantitative fashion by expert radiologist readers using validated scoring instruments. While cross-sectional reliability results between two trained and calibrated readers has been presented for all MRI scoring systems, data on longitudinal reliability in regard to detection of change over time has not been presented to date. In order to facilitate and accelerate assessment more than two radiologists may be assessing MRI datasets simultaneously. For meaningful data interpretation it is paramount to ensure reliability between all readers.
Aim of this study was to determine reliability between four different readers in cross-sectional and longitudinal fashion in the MOST study using the modified whole organ magnetic resonance imaging score (WORMS).
**Methods:** The Multicenter Osteoarthritis (MOST) study is a longitudinal cohort study of subjects with or at high risk of knee OA. 10 subjects were

**Table1. Cross-sectional and longitudinal reliability for substudy A (60 and 84 months readings, n = 10)**

| MRI Feature | Category | R01 vs. R02 | R01 vs. R03 | R01 vs. R04 | R02 vs. R03 | R02 vs. RCU | R03 vs. R04 |
|---|---|---|---|---|---|---|---|
| Cartilage morphology | Cross-sectional | 0.85 (0.81–0.90) | 0.86 (0.82–0.90) | 0.82 (0.77–0.87) | 0.77 (0.71–0.82) | 0.79 (0.73–0.84) | 0.87 (0.82–0.91) |
| | Longitudinal | 0.78 (0.67–0.90) | 0.77 (0.66–0.88) | 0.63 (0.49–0.77) | 0.63 (0.46–0.77) | 0.62 (0.46–0.77) | 0.70 (0.56–0.84) |
| Osteophytes | Cross-sectional | 0.64 (0.57–0.72) | 0.52 (0.46–0.59) | 0.47 (0.40–0.54) | 0.49 (0.42–0.56) | 0.48 (0.41–0.55) | 0.80 (0.76–0.84) |
| | Longitudinal | 0.61 (0.47–0.75) | 0.58 (0.43–0.72) | 0.54 (0.40–0.69) | 0.48 (0.33–0.64) | 0.43 (0.27–0.58) | 0.76 (0.66–0.86) |
| Bone marrow lesion | Cross-sectional | 0.89 (0.84–0.94) | 0.81 (0.74–0.88) | 0.81 (0.74–0.88) | 0.80 (0.74–0.87) | 0.83 (0.76–0.90) | 0.86 (0.81–0.92) |
| | Longitudinal | 0.88 (0.80–0.96) | 0.82 (0.73–0.92) | 0.80 (0.70–0.91) | 0.75 (0.63–0.87) | 0.79 (0.68–0.91) | 0.80 (0.69–0.91) |
| Subchondral cysts | Cross-sectional | 0.68 (0.46–0.90) | 0.54 (0.32–0.77) | 0.50 (0.27–0.72) | 0.51 (0.29–0.73) | 0.48 (0.26–0.69) | 0.93 (0.82–1.00) |
| | Longitudinal | 0.60 (0.23–0.97) | 0.70 (0.39–1.00) | 0.70 (0.39–1.00) | 0.60 (0.29–0.91) | 0.60 (0.29–0.91) | 1.00 (1.03–1.00) |
| Bone attrition | Cross-sectional | 0.79 (0.70–0.89) | 0.76 (0.69–0.83) | 0.71 (0.53–0.80) | 0.79 (0.70–0.88) | 0.80 (0.70–0.89) | 0.88 (0.80–0.95) |
| | Longitudinal | 0.71 (0.48–0.95) | 0.67 (0.46–0.88) | 0.61 (0.88–0.83) | 0.51 (0.24–0.78) | 0.55 (0.28–0.82) | 0.77 (0.58–0.97) |
| Meniscal tears | Cross-sectional | 0.94 (0.90–0.98) | 0.92 (0.88–0.97) | 0.92 (0.87–0.97) | 0.97 (0.93–1.00) | 0.92 (0.87–0.98) | 0.96 (0.92–0.99) |
| | Longitudinal | 0.92 (0.81–1.00) | 0.84 (0.68–1.00) | 0.75 (0.55–0.95) | 0.91 (0.79–1.00) | 0.81 (0.63–1.00) | 0.89 (0.73–1.00) |
| Menisci extrusion | Cross-sectional | 0.83 (0.69–0.98) | 0.82 (0.67–0.97) | 0.86 (0.72–1.00) | 0.71 (0.52–0.90) | 0.67 (0.45–0.88) | 0.81 (0.65–0.97) |
| | Longitudinal | 0.81 (0.62–1.00) | 0.75 (0.55–0.95) | 0.95 (0.87–1.00) | 0.67 (0.42–0.91) | 0.77 (0.57–0 96) | 0.81 (0.60–1 00) |
| Hoffa–synovitis | Cross-sectional | 0.60 (0.38–0.83) | 0.58 (0.36–0.80) | 0.45 (0.24–0.66) | 0.16 (0.10–0.42) | 0.24 (–0.04–0.52) | 0.59 (0.350.82) |
| | Longitudinal | 0.64 (0.00–1.00) | 0.44 (–0.21–1.00) | 0.31 (0.27–0.90) | 0.64 (0.00–1.00) | 0.45 (–0.15–1.00) | 0.77 (0.35–1.00) |
| Effusion-synovitis | Cross-sectional | 0.89 (0.75–1.00) | 0.88 (0.72–1.00) | 0.72 (051–0.92) | 0.78 (0.56–0.99) | 0.62 (0.40–0.84) | 0.57 (0.32–0.82) |
| | Longitudinal | 0.85 (0.57–1.00) | 0.64 (0.28–1.00) | 0.85 (0.57–1.00) | 0.47 (0.08–0.86) | 0.70 (0.28–1.00) | 0.47 (0.08–0.86) |

**Table 2**
Cross -sectional and longitudinal reliability for substudy B (baseline, 60 and 84 months readings, n = 10).

| 90% MRI Feature | Category | R01 vs. R02 | R01 vs. R03 | R01 vs. R04 | R02 vs. R03 | R02 vs. 04 | R03 vs. R04 |
|---|---|---|---|---|---|---|---|
| Cartilage morphology | Cross-sectional | 0.85 (0.81–0.89) | 0.86 (0.81–0.9) | 0.86 (0.82–0.9) | 0.96 (0.95–0.98) | 0.95 (0.93–0.97) | 0.93 (0.90–0.95) |
| | Longitudinal | 0.51 (0.41–0.61) | 0.51 (0.41–0.61) | 0.50 (0.41–0.60) | 0.82 (0.76–0.88) | 0.82 (0.75–0.89) | 0.67 (0.59–0.75) |
| Osteophytes | Cross-sectional | 0.92 (0.9–0.94) | 0.93 (0.91–0.95) | 0.93 (0.91–0.95) | 0.94 (0.92–0.96) | 0.94 (0.92–0.96) | 0.95 (0.93–0.97) |
| | Longitudinal | 0.56 (0.46–0.66) | 0.55 (0.44–0.65) | 0.49 (0.37–0.60) | 0.71 (0.62–0.8) | 0.60 (0.48–0.71) | 0.65 (0.53–0.76) |
| Bone marrow lesion | Cross-sectional | 0.86 (0.82–0.91) | 0.85 (0.80–0.91) | 0.93 (0.89–0.97) | 0.89 (0.84–0.94) | 0.85 (0.80–0.90) | 0.86 (0.82–0.91) |
| | Longitudinal | 0.8 (0.73–0.87) | 0.71 (0.63–0.80) | 0.88 (0.82–0.94) | 0.77 (0.68–0.85) | 0.76 (0.68–0.84) | 0.80 (0.73–0.87) |
| Subchondral cysts | Cross-sectional | 0.83 (0.69–0.96) | 0.83 (0.69–0.96) | 0.80(0.66–0.94) | 1.00(1.00–1.00) | 0.88 (0.77–0.98) | 0.88 (0.77–0.98) |
| | Longitudinal | 0.00 (0.00–0.00) | 0.00 (0.00–0.00) | 0.00 (0.00–0.00) | 1.00 (1.00–1.00) | 0.50 (0.07–0.92) | 0.50 (0.07–0.92) |
| Bone attrifion | Cross-sectional | 0.87 (0.77–0.96) | 0.94 (0.88–1.00) | 0.96(0.92–1.00) | 0.92(0.84–1.00) | 0.87 (0.78–0.96) | 0.91 (0.83–0.98) |
| | Longitudinal | 0.66 (0.35–0.97) | 0.80 (0.52–1.00) | 0.66 (0.35–0.97) | 0.56 (0.25–0.88) | 0.75 (0.50–0.99) | 0.56 (0.25–0.88) |
| Meniscal tears | Cross-sectional | 0.94 (0.90–0.98) | 0.95 (0.91–0.99) | 0.89 (0.83–0.95) | 0.97 (0.94–0.99) | 0.95 (0.90–0.99) | 0.91 (0.86–0.97) |
| | Longitudinal | 0.80 (0.64–0.95) | 0.80 (0.64–0.96) | 0.56 (0.35–0.76) | 0.92 (0.83–1.00) | 0.79 (0.65–0.93) | 0.70 (0.55–0.85) |
| Meniscal extrusion | Cross-sectional | 0.79 (0.64–0.93) | 0.75 (0.60–0.90) | 0.72 (0.57–0.87) | 0.9 (0.80–1.00) | 0.93 (0.84–1.00) | 0.89 (0.79–1.00) |
| | Longitudinal | 0.42 (0.09–0.75) | 0.13 (–0.17–0.44) | 0.19 (–0.14–0.53) | 0.57 (0.26–0.88) | 0.84 (0.62–1.00) | 0.52 (0.18–0.85) |
| Hoffa– synovitis | Cross-sectional | 0.76 (0.62–0.9) | 0.76 (0.62–0.90) | 0.73 (0.58–0.88) | 1.00 (1.00–1.00) | 0.97 (0.91–1.00) | 0.97 (0.91–1.00) |
| | Longitudinal | 0.39 (0.12–0.66) | 0.39 (0.12–0.66) | 0.21 (–0.03–0.45) | 1.00 (1.00–1.00) | 0.65 (0.21–1.00) | 0.65 (0.20–1.00) |
| Effusion–synovitis | Cross-sectional | 0.76 (0.60–0.93) | 0.71(0.55–0.88) | 0.66 (0.46–0.87) | 0.88 (0.73–1.00) | 0.96 (0.88–1.00) | 0.89 (0.75–1.00) |
| | Longitudinal | 0.51 (0.28–0.75) | 0.54(0.31–0.77) | 0.37 (0.11–0.64) | 0.70 (0.36–1.00) | 0.82 (0.57–1.00) | 0.59 (0.20–0.97) |

randomly included in this substudy (=A) that had 60 months and 84 months MRIs available. Another 10 participants were included with baseline, 60 and 84 months MRIs (=substudy B). MRI was performed at a 1.0 T extremity system using axial and sagittal proton-density weighted sequences and a coronal STIR sequence. MRIs were read by four radiologists separately with the chronological sequence known to the readers. For substudy B, readers were aware of the baseline scores. MRIs were assessed semiquantitatively using a modified WORMS system. Assessed were cartilage, osteophytes, bone marrow lesions, subchondral cysts, bone attrition, meniscus damage, meniscal extrusion, Hoffa-synovitis, effusion-synovitis, cruciate and collateral ligaments, popliteal cysts, tibio-fibular cysts, loose intra-articular bodies and anserine and pre-patellar bursitis. Weighted kappa statistics were applied to determine reliability between readers (R1 vs. R2, R1 vs. R3, R1 vs. R4, R2 vs. R3, R2 vs. R4, R3 vs. R4) for cross-sectional (for 60 months only) and longitudinal (i.e. evaluation of change) assessment.
**Results:** Subjects were on average 65.4 years old (SD ± 7.4) with 12 (60%) women and mean BMI of 29.8 (SD ± 5.0). Baseline Kellgren-Lawrence grades were 0 for 2 knees, grade 1 for 7 knees, grade 2 for 6 knees and grade 3 for 5 knees. For substudy A, cross-sectional reliability of cartilage scoring ranged between 0.85 and 0.96 and longitudinal between 0.50 and 0.82. BML assessment showed w kappa values between 0.85 and 0.93 (cross sectional) and 0.71 and 0.88 (longitudinal). Meniscal tears were scored between 0.89 and 0.97 (cross-sectional) and 0.56 and 0.80 (longitudinal). Osteophytes were assessed with kappas ranging between 0.92 and 0.95 (cross-sectional) and 0.49 and 0.71. Most other osteoarthritis features were evaluated with kappa values between 0.5 and 1.0 (Table 1). Results for substudy B were comparable in regard to agreement (Table 2).
**Conclusions:** Semiquantitative OA assessment on MRI shows good reliability for up to four trained and calibrated readers. Cross-sectional reliability seems to be slightly superior compared to scoring of change. Reliability did not differ for readings of three time points with baseline known to the readers or two time points without knowledge of baseline scores. Some of the scores at both ends of the reliability scale have to be interpreted in light of low prevalence of some of the features.

## 465
## DEFINING RADIOGRAPHIC KNEE OSTEOARTHRITIS: A COMPARISON BETWEEN THE KELLGREN & LAWRENCE CLASSIFICATION AND OARSI ATLAS

A. Culvenor †, C. Engen ‡, B. Øiestad §, L. Engebretsen §, M. Risberg ‡,§.
† *Univ. of Queensland, Brisbane, Australia;* ‡ *Norwegian Sch. of Sports Sci., Oslo, Norway;* § *Oslo Univ. Hosp., Oslo, Norway*

**Purpose:** Many different classification systems exist to define the presence of radiographic knee osteoarthritis (OA). Each system has different cut-offs to define the presence of radiographic knee OA, making comparisons of OA prevalence across studies problematic. Two of the most widely used radiographic classification systems are the Kellgren & Lawrence (KL) system and the Osteoarthritis Research Society International (OARSI) atlas criteria. While it has been suggested that the cut-off for defining knee OA is similar for both systems, no direct comparison has been undertaken. Thus, the aims of this study were to investigate and compare radiographic tibiofemoral OA rates using the KL system and the OARSI atlas criteria, compare qualitative (KL and OARSI) and quantitative (millimetres) measures of joint space narrowing (JSN) and evaluate reliability of the two systems.
**Methods:** Six-hundred and twenty-one individuals (1,242 knees) aged 40-80 years with self-reported hand, hip or knee OA were recruited from a population-based cohort study in Norway. Standardised posteroanterior radiographs (using a Synaflex frame) of 1,178 knees free of arthroplasty or osteotomy were graded with the KL system (grade 0–4) and OARSI atlas. KL grade 2 was separated into KL2/ost (osteophyte alone) or KL2+ (ostephyte and possible JSN). The OARSI atlas includes both osteophytes and JSN, individually scored from grade 0 (no change) to grade 3 (severe change). Minimum joint space width (mJSW) was measured manually in millimetres with digital callipers. All evaluations were performed by a single physician who was trained by an orthopaedic surgeon with over 20 years of experience in radiographic assessment. Intra- and inter-rater reliability was evaluated from 60 radiographs.
**Results:** According to the KL system (≥KL2+), 167 knees (14.2%) had tibiofemoral OA and 203 (17.3%) knees had ≥KL2/ost. In contrast, 309 knees (26.2%) had tibiofemoral OA according to OARSI atlas criteria (osteophyte or JSN grade ≥ 2 or grade 1 JSN in combination with a grade 1 osteophyte). KL and OARSI JSN descriptions were significantly associated with mJSW measured quantitatively ($p < 0.022$). Intra-rater and inter-rater reliability for the KL system (weighted-kappa: 0.97 (0.92–1.0) and 0.67 (0.54–0.79), respectively) and OARSI atlas (weighted-kappa: 0.86 (0.78–0.95) and 0.73 (0.63–0.84), respectively) was acceptable.
**Conclusions:** Radiographic tibiofemoral OA was almost twice as common using the OARSI atlas criteria compared to using the KL system. This discrepancy is likely to contribute to the large variability of OA prevalence observed in the literature. The cut-off for defining radiographic tibiofemoral OA with the OARSI atlas should no longer be considered equivalent to the KL system cut-off of ≥grade 2, or to the proposed cut-off incorporating a definite osteophyte alone. Future studies should include MRI to examine which of the radiographic classification systems most accurately represent radiographic structural joint changes. Consensus regarding radiographic cut-offs to define OA is urgently needed as radiography remains the recommended modality to evaluate OA presence and progression.

## 466
## DIAGNOSTIC PERFORMANCE OF WEIGHT BEARING 3D IMAGING FOR DETECTION OF KNEE OSTEOARTHRITIS FEATURES

N.A. Segal †, J.A. Lynch ‡, J. Niu §, A. Guermazi ‖, M.C. Nevitt ‡, J.C. Torner †. † *Univ. of Iowa, Iowa City, IA;* ‡ *Univ. of California, San Francisco, San Francisco, CA;* § *Boston Univ. Sch. of Med., Boston, MA;* ‖ *Boston Univ. Med. Ctr., Boston, MA*

**Purpose:** To evaluate the sensitivity, specificity, accuracy and predictive value of weight-bearing computerized tomography (WBCT) of the knee in depicting osteophytes and subchondral cysts, in comparison with fixed-flexed radiographs, using magnetic resonance imaging (MRI) as the reference standard.
**Methods:** Out of 83 MOST participants, who met inclusion criteria (had knee radiographs and MRI in the past year that met MOST reading center quality control criteria, live in proximity to Iowa City and distal thigh width did not exceed the 38.1 cm gantry), 19 were recruited with knees discordant for KL grade in order to represent a range of OA features. Following informed consent, participants stood on the scanner platform. A custom positioning system was used to maintain foot external rotation and fixed knee flexion angles, with participants' thighs and hands contacting the unit for stability and prevention of motion. The scanner produced pulsed cone-beam x-ray (effective dose equivalent 10 mrem) on a 30 × 30 cm amorphous silicon flat-panel detector over a 360° projection angle with total scan time of 32 seconds. A 3D axial CT dataset with isotropic resolution of 0.37 mm and FOV of 350 mm was reconstructed from initial cone-beam projection images. Marginal osteophytes and subchondral cysts were scored on bilateral radiographs and bilateral WBCT images (axial reconstructed as coronal plane images) according to the OARSI grading system and dichotomized into either present (grade 1 or above) or absent (grade 0). For WBCT, the reader selected the coronal slice that best demonstrated the joint space and osteophytes. All images were read in a random order, and radiographs and WBCT images were read during separate sessions, with a 2-week interval, to prevent simultaneous assessment of both modalities, which could potentially bias the readings. MRIs (axial and sagittal proton density-weighted and coronal STIR sequences) of the same knees were also read for the same features, also with a 4-week interval from reading other modalities. The presence of marginal osteophytes and subchondral cysts detected with conventional radiography and WBCT were compared with those detected with MRI of the knee, as a whole and according to location (i.e., medial or lateral, femur or tibia). MRI findings were used as the reference standard. Sensitivity, specificity, and accuracy for the detection of those lesions were calculated in the left and right knees separately and McNemar's test was used to examine for statistically significant differences between imaging modalities. Positive and negative predictive values also were calculated, and using generalized estimating equation marginal regression models, we assessed the significance of the differences.
**Results:** Of the 19 participants, there were 8 women and 9 participants were age 50–59, whereas 10 were age 60–69. The mean BMI was 30.0 kg/m$^2$. Of the 152 surfaces (medial and lateral femur and tibia for 38 knees), MRI revealed 79 osteophytes and 10 subchondral cysts (Table 1). In comparison with osteophytes and subchondral cysts detected by MRI, WBCT was significantly more sensitive and accurate than plain radiographs (Table 2). For osteophytes, the greatest differences in sensitivity and accuracy were detected at the medial femur. A low rate of subchondral cysts as well as a low rate of false positive osteophytes on